GMS 4.0: New Modeling Tools For Stratigraphic and Stochastic Modeling and Uncertainty Analysis

Cary A. Talbot[1], Clarissa M. Hansen[2], Dr. Norman L. Jones[3] and Earl V. Edris[4]

[1] Research Hydraulic Engineer, CEERD-HC-HG, 3909 Halls Ferry Road, Vicksburg, MS 39180, 601-634-2625, Cary.A.Talbot@erdc.usace.army.mil

[2] Research Hydraulic Engineer, CEERD-HC-HG, 3909 Halls Ferry Road, Vicksburg, MS 39180, 601-634-2102, Clarissa.M.Hansen@erdc.usace.army.mil

[3] Director, Environmental Modeling Research Laboratory, 242C Clyde Building, Brigham Young University, Provo, UT 84602, 801-422-2812, njones@et.byu.edu

[4] Chief, Hydrologic Systems Branch, CEERD-HC-H, 3909 Halls Ferry Road, Vicksburg, MS 39180, 601-634-3378, Earl.V.Edris@erdc.usace.army.mil

Abstract

With the introduction of version 4.0 of the Department of Defense Groundwater Modeling System (GMS), users are provided with several new tools that will greatly enhance the ability to model complex sites and deal with the inherent uncertainty of subsurface systems. A key component to successful modeling of subsurface hydrogeologic systems is the development of a suitably accurate representation of the stratigraphic system present at the site. Past methods for building such representations have often required over-simplification of stratigraphy in order to build models using the available tools. Stratigraphy models built using the new Horizons Method fielded in GMS 4.0 allow users the ability to include heterogeneity and complex geologic conditions in their models. Another often-encountered problem in groundwater modeling is the inherent uncertainty associated with parameters such as the hydraulic conductivity of a geologic material. Stochastic modeling tools as well as tools for generating probabilistic realizations of geologic material distributions have been added to GMS 4.0. These tools allow the user to present groundwater modeling results in probabilitistic fashion. New tools for performing uncertainty analyses are also provided which can be used as input for risk analysis studies. Because stochastic modeling often generates a large number of model results, a new tool called the Data Tree is included which is used for organizing all data in GMS 4.0. This paper describes these new enhancements to GMS and offers examples of how they can be used to help users address the dual challenges of uncertainty in the subsurface and conducting groundwater modeling studies at sites with complex hydrogeologic conditions.

Introduction

A suite of multidimensional tools for performing computational studies of hydrologic, hydraulic and geohydrologic systems has been developed and fielded by the US Army Engineer Research and Development Center (ERDC) for use within the US Army Corps of Engineers (Corps), other government agencies and the commercial sector. These tools consist of modeling systems that provide a single environment for users to carry out all phases of a "hydroinformatic" study. The term hydroinformatic is often used to describe the combination of computational, data processing and decision-making tasks that must be performed in carrying out hydraulic, hydrologic and geohydrologic studies. The Department of Defense (DoD) Groundwater Modeling System (GMS) is the subsurface component of the Corps' multidimensional hydroinformatic toolbox.

GMS first began development in the 1990's after it had become apparent that available tools for performing groundwater modeling studies were ill suited to meet the challenges of scale, complexity and scope that were being identified at military installations and in civil works projects within the Corps' and DoD's realm of responsibility (Richards et al. 1998). Since the release of GMS version 1.0 in November of 1994, numerous tools have been developed and added to GMS that enhance its ability to provide users with state-of-the-art capabilities and numerical modeling codes. These tools have been released to the GMS user community in the form of new version releases and updates to existing versions. With the release of GMS version 4.0 in October 2002, several significant additions have been made to the system that provide users with improved tools for data organization and stratigraphic modeling as well as new tools for performing stochastic modeling and uncertainty analysis.

Stratigraphic Modeling

The earliest versions of GMS included tools for developing digital three-dimensional "solid models" of the subsurface stratigraphy. A solid model is defined from borehole data and surfaces that represent the boundaries between unique stratigraphic units or materials. A solid model completely and unambiguously defines the volumes of three-dimensional subsurface materials. Pinchouts, embedded seams and faults can all be directly represented in the solid model geometry with no voids or overlaps between the boundaries of the individual materials if the solids are defined properly (Lemon & Jones 2001). However, generating such solids can be a time-consuming and difficult task, particularly with complex stratigraphy and numerous layers of materials.

Horizons Method. GMS 4.0 introduces a new and simpler method for generating solids from borehole data that seeks to address the problems associated with generating solid models of geologically complex sites. This method is called the *horizons method* and is based on the notion that the top of each stratigraphic unit that will be represented in the solid model can be considered a "horizon" (Lemon & Jones 2001). Horizons are numbered consecutively according to their depositional order (i.e. from the bottom up) and assigned at the interfaces between materials in each borehole (the borehole "contacts"). Each contact of each material that is to be represented in the solid model must have a horizon id. Figure 1 depicts a set of boreholes with horizon ids assigned. Any borehole contact the user wishes to ignore in the construction of the solid model can be assigned a horizon id of zero.

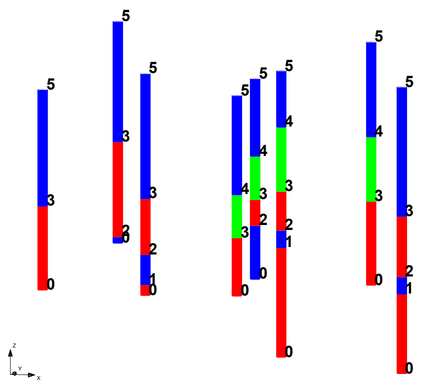

Figure 1. Boreholes with horizon ids assigned.

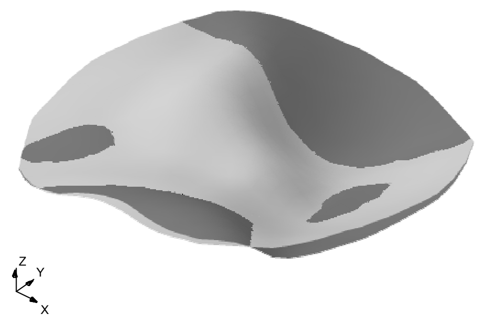

Figure 2. Horizon surfaces created by interpolating from borehole contacts.

Horizons->Solids. Solid models are constructed from boreholes with assigned horizon ids using the *Horizons->Solids* command. The process of creating the solids occurs automatically and rapidly without any further user intervention, however it is instructive to discuss the process by which these solids are constructed by GMS. First, all contacts assigned horizon ids greater than zero are converted to scatter points with a data set for each horizon id. A surface is interpolated from each data set using a user-selected higher-order interpolation scheme. The resulting surfaces represent each material horizon. Figure 2 depicts two horizon surfaces generated from such an interpolation, the dark grey surface corresponding to horizon id 1 and the light grey surface corresponding to horizon id 2. A solid is first created by extruding the horizon id 1 surface down to a surface defined by the bottom of the boreholes. The solid for horizon id 2 is then created by filling in the areas where the horizon id 2 surface is above that of horizon id 1. This sequence is repeated for all materials with the new solid being created by subtracting the existing solids from the solid extruded from the next higher horizon surface (Lemon & Jones 2001). A set of solids created by the horizons method using the boreholes and horizon id assignments found in Figure 1 is shown in Figure 3.

Borehole Cross-sections. The horizons method can greatly accelerate the construction of solid models from borehole data. However, the accuracy of the resulting solids is highly dependent on the quality and quantity of borehole data available. Often, due to the distance between borehole locations, pinchouts, seams and other complex features cannot be appropriately defined based on borehole data alone. Some degree of geologic interpretation is needed to construct solids that depict an estimated or accepted stratigraphic configuration in greater detail. GMS 4.0

also includes a tool for constructing cross sections between any two boreholes that allow the user to define the configuration of the geologic materials in the areas between boreholes. As shown in Figure 4, the user can define the shape of the horizon surfaces in the space between boreholes by using piece-wise linear arcs. These arcs comprise a user-defined cross-section that is used to define the shape of the horizon surfaces between boreholes and the resulting solids created from the surfaces. User-defined borehole cross-sections are a powerful tool for generating detailed depictions of highly varied and complex stratigraphy as demonstrated by the solid model built for the Department of Energy Paducah Gaseous Diffusion Plant study (Figure 5). A total of 120 boreholes and 31 user-defined borehole cross-sections were used to construct this solid model of highly varied and complex stratigraphy.

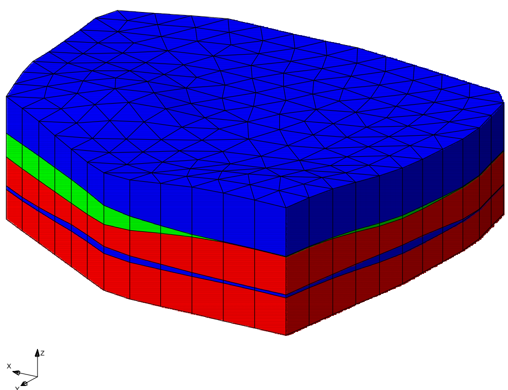

Figure 3. Solid model created by the horizons method.

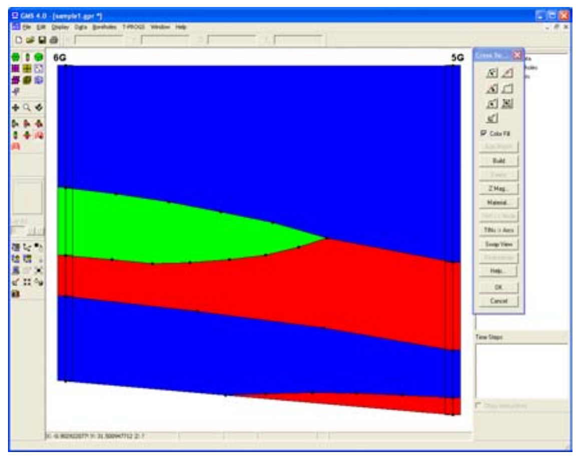

Figure 4. Borehole cross-section editor in GMS 4.0.

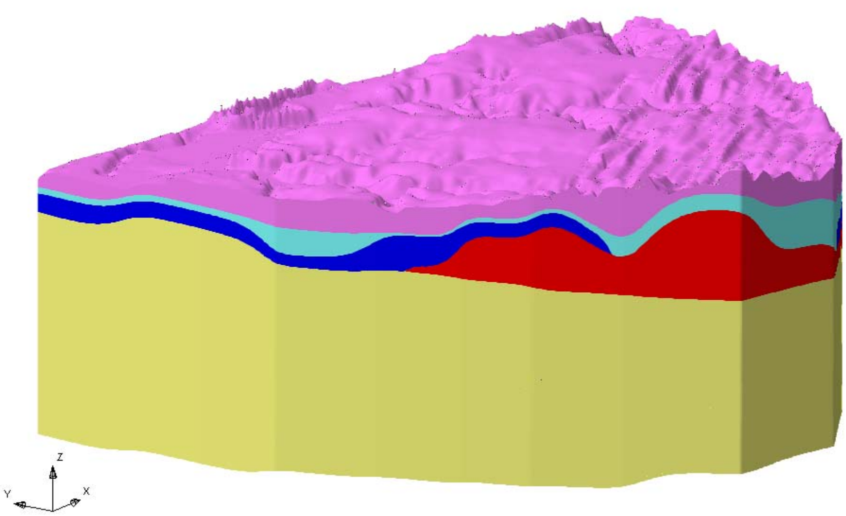

Figure 5. Paducah Gaseous Diffusion Plant solid model.

Stochastic Modeling Tools

The process of numerical modeling has long consisted of using numerical techniques to solve boundary-value problems in a deterministic fashion. Unknown values on the interior of a boundary are found by determining a set of parameters for which the governing equations can be solved to within an acceptably small error tolerance. This approach, however, assumes that there is a *unique* set of parameters that will meet the solution criteria. In the realm of subsurface modeling, the uncertainty of parameter estimates based on field data, variability of subsurface stratigraphic conceptualizations and the applicable numerical model assumptions render it entirely possible that multiple parameter sets could offer equally valid solutions. As an alternative to deterministic approaches, stochastic modeling, where equally probable sets of parameters are used to generate a range of solutions with associated probabilities, are currently being applied to subsurface modeling problems at an increasing rate. Two stochastic modeling tools have been introduced in GMS 4.0. The first allows the user to randomize selected model parameters within specified ranges and the second generates multiple realizations of the distribution of subsurface materials with parameters being assigned to the model domain based on material type. Post-processing tools have also been introduced in GMS 4.0 to provide automated generation of two types of analysis common to stochastic modeling: probabilistic capture zone analysis and probabilistic threshold concentration analysis.

Parameter Randomization. With the parameter randomization approach, the user first defines zones where a particular parameter value is uniformly assigned to a portion of the model domain, just as is frequently done in deterministic modeling approaches. However, ranges of equally probable values are assigned to each parameter zone based on the uncertainty associated with that particular parameter. MODFLOW2000, which is supported in GMS 4.0, includes automated tools for assigning these parameter ranges. A user-specified number of sets of parameter values are then determined in random fashion using either the Monte Carlo or Latin Hypercube methods (Figure 6). Not all simulations will converge due to the random nature of the parameter values therefore automated tools are provided in GMS for organizing the simulation results based on convergence and goodness of fit with observed data. The greater number of simulations that are run increases the

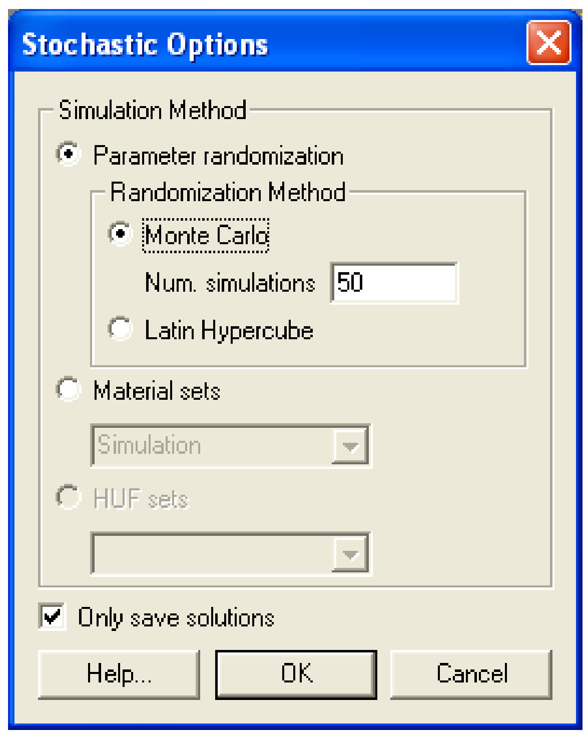

Figure 6. Stochastic Modeling Options in GMS 4.0.

confidence that the parameter space has been adequately explored (Green & Jones 2001).

Material Sets. Another approach to stochastic modeling is to associate a set of parameter values with a particular material and then generate multiple realizations of material distribution (or "set") within the model domain. Two geostatistical techniques are provided in GMS 4.0 to generate a user-specified number of equal probability material sets: indicator kriging, based on the UNCERT code developed at the Colorado School of Mines (Wingle et al.1999) and multi-dimensional Markov chains using Carle's T-PROGS software suite (Jones et al. 2002; Carle & Fogg 1997). Both of these techniques generate material sets conditioned on available hard data such as boreholes or scatter points. In the case of the T-PROGS approach, vertical borehole data are analyzed to generate a matrix of probability distribution curves. These curves include data defining the proportion of each material, the average dimension of the material lenses, and juxtapostioning relationships. Once the vertical data are analyzed, horizontal curves (Markov chains) are generated by a combination of data extracted from the vertical curves and some additional data provided by the user. The multi-dimensional Markov chains are then transformed to generate indicator kriging equations (Jones et al. 2002). Each material set generated by either approach represents an equally probable realization of the subsurface distribution of materials (Figure 7). Parameters are assigned based on material type and simulations for each material set are then run. The resulting set of simulation results represent a set of equally probable solutions based on the randomization of materials.

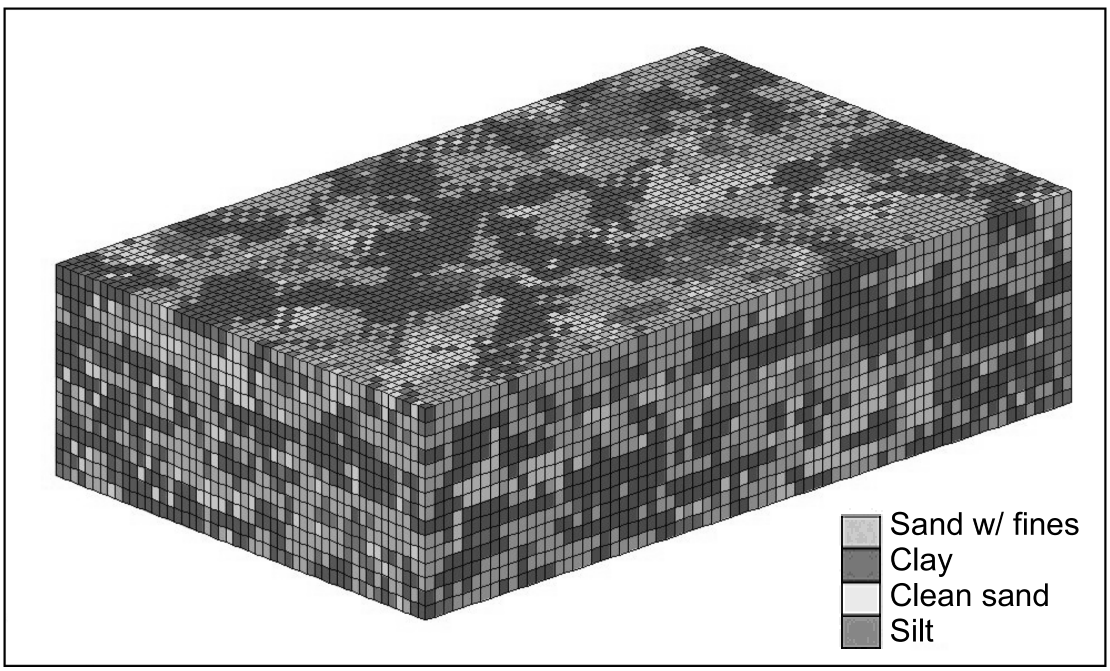

Figure 7. Material set generated from T-PROGS approach.

Concentration Threshold Analysis. When a stochastic approach is used to generate multiple realizations of a groundwater flow solution, those flow fields can then be used to generate an equal number of contaminant transport concentration solutions. GMS 4.0 provides a tool for performing automated concentration threshold analysis based on these solutions. To perform such an analysis, the user starts with the solutions loaded in to GMS and then specifies a concentration value, or threshold, that is of particular interest. A concentration probability data set is created with one value for every cell of the model grid, the values all being zero initially. GMS then examines the set of solutions and for each solution, compares the concentration value in each cell of the grid to the specified threshold. If a cell is found to have a concentration greater than the threshold, a value of (1/n), where n is the total number of simulations in the set, is added to the corresponding value in the concentration probability data set. The resulting data set consists of percentage values for each cell of the model grid indicating the probability that the threshold concentration is exceeded by the solution set values in that cell. These percentages can then be contoured to generate a threshold exceedance map like that shown in Figure 8.

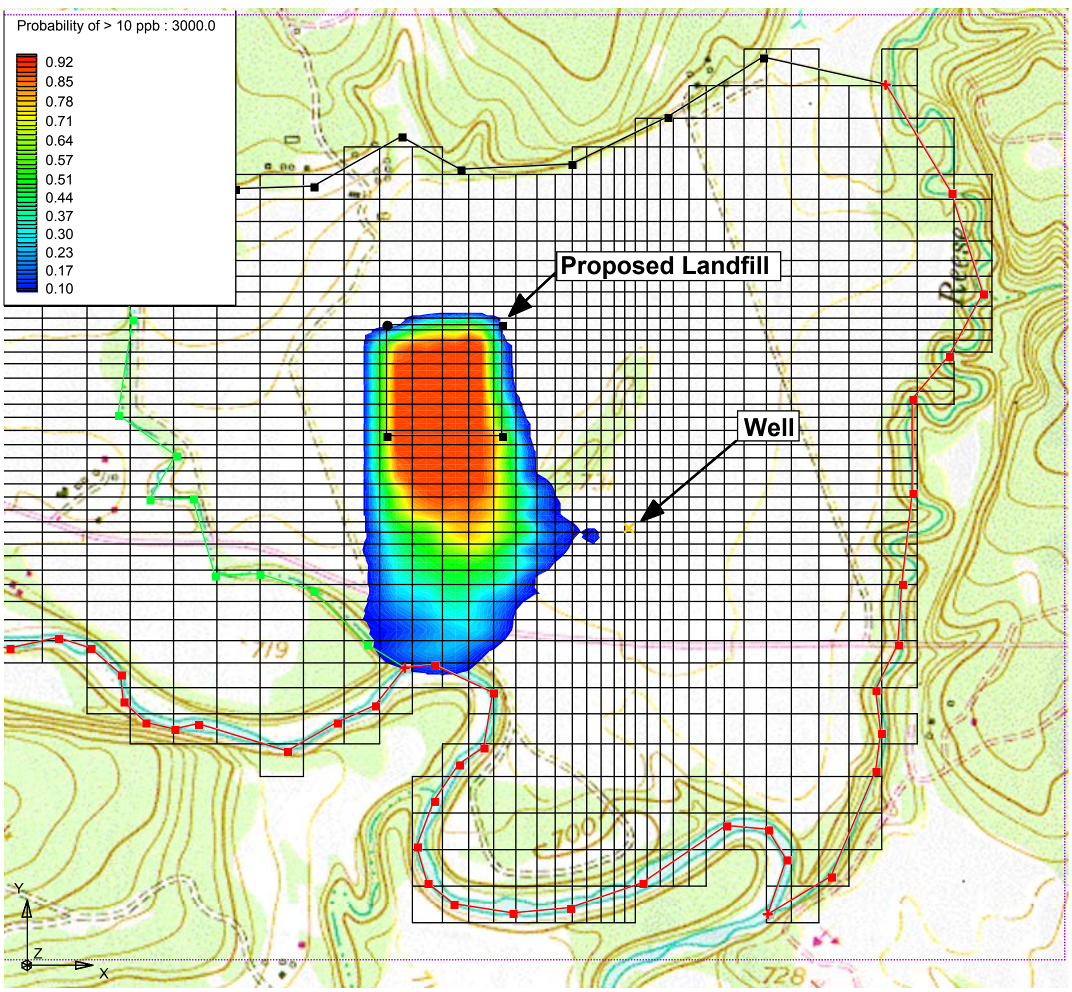

Figure 8. Threshold exceedance map. Contours are of the probability of exceeding 10 ppb.

Capture Zone Analysis. Capture zone analysis is an often-used analysis tool in groundwater simulations, particularly wellhead protection studies. However, when a stochastic modeling simulation has been performed, the set of simulation solution files can be used to generate a capture zone risk map that contours the probability of capture of any given well. GMS 4.0 provides an automated tool for performing this type of probabilistic analysis. As with the concentration threshold analysis, a grid data set, which will store capture frequency values, is generated with all values initialized to zero. Particles are then placed in each cell and tracked forward in time using the MODPATH code. If the particle is captured by a well, the value in the capture frequency data set corresponding to the starting cell of that particle is incremented. After all simulations in the solution set have been tested, the frequency array is then divided by the total number of solutions, giving a data set of percentage of capture probability in each cell of the grid. If a simulation contains more than one well, a separate capture frequency data set is created for each well. Frequency capture data sets can be used to generate contoured capture zone risk maps like that shown in Figure 9.

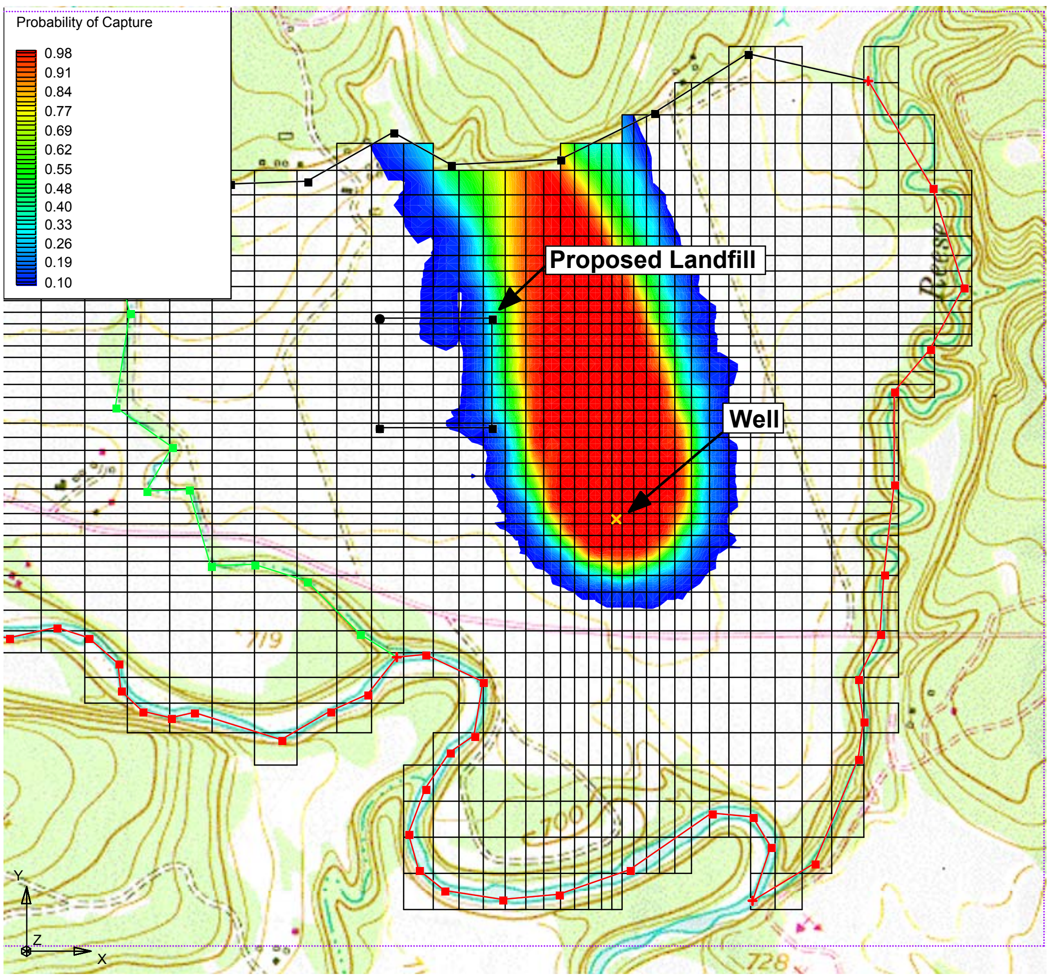

Figure 9. Capture zone analysis. Contours are of probability of capture of flow by the well.

Data Tree

With the addition of stochastic modeling tools in GMS 4.0, the number of model solution and other data sets that can be generated in GMS increased dramatically. The data set organization tool used in previous versions of GMS, the Data Browser, was inefficient at handling such a large number of data sets in memory at the same time. As well, because a user may wish to, for example, rapidly view a particular set of display options applied to all 50 solutions of a stochastic model simulation, the persistence presence of a data set organization tool on the GMS desktop was needed. The Data Tree fielded in GMS 4.0 is designed to facilitate the handling of data sets in GMS, whether there be many or few, by using a hierarchical folders and files tree-like display that should be familiar to any user of Microsoft Windows. The data tree is, by default, displayed on the right-most side of the GMS desktop at launch and is always visible (Figure 10). The Data Tree window can be moved by selecting the bar above the Data Tree display and dragging the window to any location on the users' desktop, including outside of the main GMS window itself.

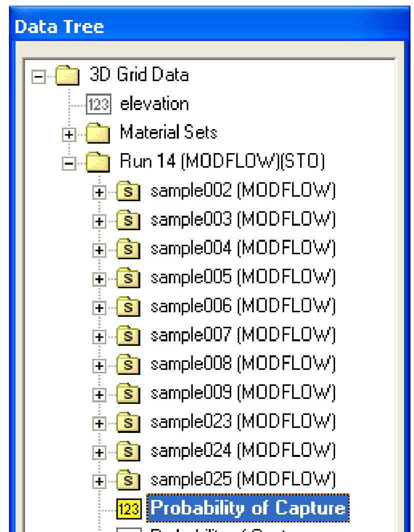

Figure 10. Data Tree window in GMS 4.0.

The contents of the Data Tree window change depending on the currently active module in GMS. For each module, the user is shown the various objects along with any currently defined data sets or model solution files currently in memory. Folder icons are used to represent groupings of objects or data sets that have some common relation or meaning. For example, if a coupled flow and transport FEMWATER simulation is run, the resulting model solution might contain data sets of pressure head, velocity, total head, concentration and boundary node flux. When read in to GMS these data sets are all grouped together in a folder labeled with the simulation name. The user is free to create their own hierarchy of folders to group data sets or objects as needed to organize data in a meaningful fashion.

Because of the hierarchical nature of the folders and objects, the tree can be expanded or collapsed as desired by the user using the familiar "+" and "-" symbols displayed in small boxes next to folder icons. Check boxes are also found between these symbols and the folder and object icons. These check boxes are the means whereby the display of objects and entire folders can be enabled or disabled. There are many more operations such as drag-and-drop of files, right-click features, copy-and-paste and other functions that can be performed using the Data Tree but which are best discovered by the user's own experience. The Data Tree is a significant organization tool that provides the user with improved efficiency, intuitive access and organizational capability of GMS objects and data sets.

Conclusion

GMS 4.0 provides users with many new capabilities that greatly enhance the ability to perform sophisticated subsurface investigations. Improved tools for building solid models of complex stratigraphy via the horizons method enable users to more accurately and efficiently build models of sites where natural complexity was previously accounted for in a less detailed fashion. The automated stochastic modeling tools using the material set and parameter randomization approaches save users tremendous amounts of effort and time in performing stochastic simulations. It is very likely that many users are being introduced to stochastic simulations for the first time because of the ease whereby these types of simulations can be performed in GMS 4.0. The addition of specialized visualization tools for probabilistic concentration threshold and capture zone analysis and the Data Tree with its hierarchical organization capabilities complete the suite of uncertainty analysis tools added to GMS 4.0. Individuals interested in learning more about GMS are invited to visit http://chl.wes.army.mil/software/gms to obtain more information or download a copy of GMS 4.0.

References

Carle, S.F., and Fogg, G.E. (1997). "Modeling spatial variability with one and multidimensional continuous-lag Markov chains," *Mathematical Geology,* v. 29, n. 7, p. 891-918.

Green, J.I., and Jones, N. L. (2001) "Tools for stochastic simulations in GMS," *Proceedings of the MODFLOW 2001 and Other Modeling Odysseys Conference,* Golden, Colorado, Sept. 11-14, 2001.

Jones, N.L., Walker, J.R., and Carle, S.F. (2002). "Using transition probability geostatistics with MODFLOW," *ModelCARE 2002: Proceedings of the Fourth International Conference on Calibration and Reliability in Groundwater Modeling,* Prague, Czech Republic, 17-20 June 2002, Vol. I, pp. 295-298.

Lemon, A.M., and Jones, N.L. (2001). "Managing complex stratigraphy in MODFLOW models," *Proceedings of the MODFLOW 2001 and Other Modeling Odysseys Conference,* Golden, Colorado, Sept. 11-14, 2001.

Richards, D.R., Talbot, C.A., and Howington, S.E. (1998). "Hydroinformatic issues in gaining regulatory acceptance of modeling technologies" *Proceedings of Hydroinformatics '98,* Copenhagen, Denmark, Aug. 24-26, 1998.

Wingle, W.L., Poeter, E.P., and McKenna, S.A. (1999). "UNCERT: Geostatistics, Uncertainty Analysis, and Visualization Software Applied to Groundwater Flow and Contaminant Transport Modeling," *Computers and Geosciences,* 25(4), pp. 365-376.